Bull. Nov. Comp. Center, Comp. Science, 45 (2021), 15-28-
@ 2021 NCC Publisher

Research and forecasting of passenger traffic using the MIX-PROSTOR system

Mikhail Bulyonkov, Vladimir Malov, Tayana Nesterenko

Abstract The paper considers the problem of evaluating the possibilities of alleviating the transportation discrimination of the population of Asian Russia. The authors suggest extending the MIX-PROSTOR system for transportation stream modeling by including tools for evaluating the options of multimodal passenger transportation.

Keywords: transportation problem, modeling, research automation system

Introduction

Recent years saw an increase in attention to the problem of population reduction in Asian Russia, and in its northern regions in particular. Moreover, the recently announced Russian Ark project, based on the concept of creating new cities in Siberia and the Far East, is aimed at implementing the concept of outrunning growth of the country's eastern regions. The total population of these news cities is estimated at over 1 million people. The current trend demonstrates the opposite: a steady decline of the population of the eastern regions of the country [1, 2]. At the same time, the whole population of the country is failing to grow at rates typical for the last century.

There is a gradual understanding that even the simple task of retaining the population numbers in Asian Russia is impossible without a significant and preemptive increase in the quality of life, compared to the regions of the European part of Russia. This is especially true for new employment opportunities in the new cities[1]. Even if we assume that the "new" cities are the result of a radical modernization of the existing cities, a radical improvement of the quality of life is needed. This is likely to be the motivation behind the opinions that expats are expected to return to the Eastern regions of Russia. To implement this policy, various benefits are proposed, including relocation allowances and benefits, like it happened early in the XX century, during Stolypin's reforms aimed at the Russian colonization of Siberian territories.

In any case, reasonable proposals on the organization of a new stage of settlement of Siberia and the Far East require evaluation of conditions and possibilities of the whole living complex: organization of healthcare and education, modern housing and utilities infrastructure, commuting, etc. This paper focuses on the study of the latter – the

[1] The statement of S. K. Shoigu is quite emblematic: "During our meeting with representatives of the scientific community of the Siberian Branch of the Russian Academy of Sciences, I spoke literally of the looming need to build three to five centers of science and industry in Siberia with populations of 300,000 to 1 million people. They must become the new gravitation centers for the population of Russia and our numerous expats in the countries of CIS and beyond". See https://www.rbc.ru/politics/06/09/2021/6131fab69a79471a71a0b412

conditions necessary to provide the population of Asian Russia with transportation services in order to reduce the apparent current transportation discrimination.

With the volume (both informational and computational) of the problem in mind, we understand that “reaching” each individual settlement, especially with regard to all the variants of the foreseeable future, is impossible. We will limit the scope of our study to an aggregated representation of the concentration points of population and travel routes. Moreover, we will further simplify our problem by assuming that the very concept of “transportation accessibility” concerns primarily the ability of the citizens of Siberia and the Far East to visit the central regions of European Russia: Moscow, St. Petersburg, Sochi, etc.

The validity of such simplifications requires additional justification. Accessibility of European Russia for millions of inhabitants of major cities is one thing (we are not yet considering their travels further west). Accessibility of these major cities from smaller cities and even smaller settlements is a different issue. This does not translate into millions of people, and for the latter, further travel west may not be as important, however, the requirement to provide transportation accessibility applies equally to all Russian citizens regardless of their place of residence. For this reason, our problem is concerned only with the former part. We hypothesize that these cities host the whole population of their regions. Such simplifications predefine the requirements to the use of aggregate tariffs on certain routes and transportation segments. We are not going to account for all the peculiarities of tariff formation for different means of transportation, which generally depend on the season and travel distance (with a nonlinear relation to distance) and even from the choice of the lower or upper bunk bed in a passenger car. We will try to neutralize the effect of these nuances in the original information by using interval approach both to “input” information and in the interpretation of the results[2].

In Asian Russia, there exist the following centers of population concentration:

Moscow will represent all regions of European Russia;

- Vladivostok;
- Khabarovsk, “including” the populations of Komsomolsk-on-Amur and Blagoveshchensk;
- Yuzhno-Sakhalinsk, including Kamchatka, Chukotka and Magadan;
- Yakutsk;
- Bratsk;
- Ulan-Ude;
- Irkutsk;
- Krasnoyarsk;
- Abakan;
- Novosibirsk;
- Novokuznetsk, including Kemerovo and Mezhdurechensk;
- Barnaul.

[2] This problem is exclusively passenger-oriented, yet it is an integral part of the whole MIX-PROSTOR system, which presumes comparing the passenger and cargo aspects of evaluating the capacities of the transportation system of Asian Russia [4, 5]. I.e., eventually, we plan to evaluate the carrying capacities of specific segments of different means of transportation, the possibility to organize a multimodal transportation system for all types of transportation processes, both passenger- and cargo-oriented. However, this falls outside the scope of this paper.

Cities such as Omsk, Tyumen, Salekhard, as well as all the Ural cities, are not covered. Such aggregation is explained by our interest being primarily focused on the answer to the question whether the implementation of the quest for a significant increase in transportation accessibility will affect the passenger train load on the Trans-Siberian Railway. Unless air transport is able to provide access to all settlements of Siberia and the Far East at reasonable tariffs, people will have to resort to railway and/or highway transportation to visit the cultural and health centers of European Russia. It is implied that there have been sociological studies to define the conditions on which the residents of Siberia and the Far East are willing to stay in their respective regions for permanent residence. It is also implied that the numbers of adequately paying jobs (sufficient at least to cover transportation tariffs) are retained and even increased.

It is assumed that a large, or perhaps even the main part of all passenger traffic is directed to “aggregated Moscow”. Local passenger streams (such as Barnaul-Novosibirsk) will be accounted for after the main problem is solved, which can be formulated in the following manner: to define the most rational directions and means of transportation (considering possible transfers) in order to provide transportation accessibility to the population of the eastern regions of Russia to the cultural and recreational centers of European Russia.

The concept of “rational destinations” includes such criteria as travel costs and time en route. These two criteria are considered separately and in combination. The used criteria of “rationality” are the averaged for all passengers and means of transportation, which is, in a way, a simplification, yet, as we believe, acceptable for solving the problem of systemic representation of transportation communication options. A lot will depend on the interpretation of the generated solution. In the latter, the focus will be on the quality of the result, defining the interval of tariff changes, falling outside the limits of which may lead to changes in routes and/or means of transportation.

The main goal of this study is to evaluate the possibilities of reducing the transportation discrimination of the population of Asian Russia. Minimizing future transportation costs is another goal (provided there is a choice of future options). At the same time, the study evaluates variants of future transportation networks, i.e., coordinated development of all means of transport and creation of modern logistics throughout Asian Russia. We assume that in European Russia this system has been forming more rationally and “naturally” since a lot of work was already done in this direction.

The transportation network serves not only passengers, but also, and probably to a greater degree, it serves cargo logistics. How does one separate the limited capacities of different transportation means between cargo and passengers? The example of the Trans-Siberian Railway is a telling one: more high-speed passenger trains mean fewer possibilities to transport coal. This applies to the distances of 4,000-5,000 kilometers, not to mention the importance of railways for national defense and security.

1. Modeling

Below is what we would like to propose in the result of solving our problem with the following variables (parameters):

- volumes of passenger traffic,
- limited carrying capacity of individual stretches (shoulders) of transportation,
- tariffs for these operations.

Limits applying to a “shoulder”, i.e., a segment of a transportation line, are not applicable to air transportation if limitations connected with air traffic control support are disregarded. On the other hand, the following factors are important for air transportation:

- limited capacities of airports on arrival, departure and transit of passengers, and
- limitations of the aircraft fleet in general.

Limitations of this kind have not been considered in our previous studies [6, 7], since they do not contradict the current model conceptually, as they are also expressed as linear limitations for the same variables.

By “tariffs”, we mean transportation costs, i.e., the expenditure rate. We are not concerned with the efficiency of an individual participant of the transportation process or an individual company. Our goal is to evaluate the general (systemic, total) costs of all transportation process, regardless of who incurs these expenses. It is the expenditure rate for all elements (loading, unloading, transit, reloading, and transportation) that must be the main cost indicator of efficiency. We will also attempt to convert investment into increasing the capacity into expenditure rates. Investments into expanding capacities should also be converted into expenditure rates where possible. Costs of tickets, which are currently extremely variable, may throw us off the course of our calculations: every company tries to engage more passengers by offering a variety of promotions, bonuses, etc. While this is important for companies, taking it into account will unnecessarily complicate both the solution and the interpretation of results. Again, we will focus on the relationship of aggregate tariffs, carrying capacities and volumes to be transported. Eventually, we plan to compare the requirements set out by the industry to increase the volume of transported cargo and to improve the transportation accessibility of Asian Russia. In other words, the carrying capacities of different means and points of transportation (transit, transshipping hubs, etc.) will need to be “split” between cargo and passengers by using both the efficiency criteria (minimal combined expenditures) and requirements for providing transportation accessibility.

A few remarks on the use of various transportation means:

- pipelines – relevant exclusively for the cargo segment (oil, gas). The use of pipelines for passenger transportation remains in the realm of science fiction;
- marine – almost exclusively for export cargo traffic and occasionally for cruises (tourism and recreation);
- waterways (rivers, lakes, canals) – may be the only means of transportation in some regions, but such regions are few and the numbers of passengers are low. The total of these expenses is negligibly small;
- aviation – almost exclusively for passengers, except the military and the emergency services;
- seaports – aimed almost exclusively at export-import operations, i.e., for the cargo segment of the transportation network.

Passenger traffic, as compared with cargo traffic, has certain specific features. Migration aside, all passengers return to their departure points within the extended period considered in our problem. This means that the optimal solution would be such where no one travels anywhere. To avoid this, for each pair <departure point, destination point> we must define a specific “product”. This causes a significant increase in the number of transportation products compared to cargo traffic and, as the result, the increase of the size of the problem, which has a quadratic dependency on the number of products.

This problem can be ameliorated by theorizing that the numbers of passengers transported “there” and “back” are the same. For instance, the volume of passenger traffic of “Novosibirsk-Moscow” equals that of “Moscow-Novosibirsk”. This means we can double one of these products and exclude the other.

2. Original data

With the fact that that the bulk of information will be either extremely averaged or simply unavailable (due to state or commercial secrets, or lack of records) in mind, it is more important to establish the proportion of these data for which a variant of transportation network development provides minimized combined costs. Such results may be only indirectly useful for actual logistics businesses, as information to consider. Our applied results could be useful for the Russian Railways, Ministry of Transportation and Ministry of Economic Development acting as state structures concerned with solving problems of strategic importance. Our approach (problem) can be instrumental as an imitation basis for evaluating different variants of the development of the transportation network of Asian Russia in 2035-2050 [2]. The interval format of the input information greatly alleviates the crudeness (i.e., superficial representation) of the input data[3].

There are additional requirements to the simplification of the representation of passenger streams for different means of transportation.

1. For airline passengers, one of the key limitations is the carrying capacity of airports in destination points, equivalent to the number of airplanes to be served. The type of aircraft and its passenger capacity is also important, but for the major airports featured in this problem it is approximately the same – from 150 to 200 passengers. For this reason, knowing the number of daily flights to Western airports (Moscow, Saint-Petersburg, Sochi, etc.) allows us to estimate roughly the passenger “load”.
2. For transfers from air to railway and highway transportation (and back), carrying capacity limitations are measured by the number of passengers who can be transported along these routes. For railway transportation, a possible measure is the number of pairs of passenger trains passing along the given segment. The number of passengers in each train can be estimated based on the qualitative composition of the types of cars in passenger trains (sitting, open-plan sleeper, compartment sleeper, or first-class sleeper cars) and loads of each car depending on the season derived from expert evaluations. Limitation of highways can be derived from expert evaluation of the characteristics of passenger and private highway traffic compared to cargo highway traffic.
3. With railroad transportation, a different evaluation of loads and carrying (throughput) capacity is required. We will use the stretch of railway between Novosibirsk and Omsk as the reference sample – it is a double-track electrified segment with available data on the volumes of cargo throughput and carrying capacity (pairs of trains per day). A passenger train has fewer cars than a container- or coal-carrying cargo train, but it is also a part of the pair that needs to be let through, moreover, requiring a longer safety interval. In other words, a pair of passenger trains is at least equal to a pair of coal-carrying cargo

[3] The scenario presented below is probably more appropriate in the comments section due to the character of its input information. Moreover, some parameters are classified information (e.g., the carrying capacity of specific stretches of railways and reserves in emergency situations)

trains in terms of railway carrying capacity, even though it carries much less "useful" load, in tons. Our problem requires us to evaluate the perspective of possible transfers between different means of transportation. Here, we must assume that one passenger roughly equals 4 tons of cargo. This is done based on the analogy with an open-plan sleeper car, which carries about 50 passengers. A train of 20 cars, including compartment and first-class sleepers, carries 750 passengers. In this case, a passenger train can be treated (theoretically) as a cargo train carrying 3,000 tons of cargo (calculated as the equivalent of 50 cargo cars each carrying 60 tons of load).

In all of the above cases, we will use extremely aggregated information based on expert evaluation, which will require an application of the methods of the interval approach both to the solution methods and to the representation and interpretation of the results.

For instance, the passenger traffic of Novosibirsk – WEST can be assessed in the following manner. In 2020, 40 flights left the Novosibirsk airport (OVB) daily, each carrying 150 to 200 passengers. Assuming that the planes were fully occupied gives us 6000-8000 passengers per day, and with 350 operating days a year (assuming that the airport is closed 15 days a year), this translates into 2.1 – 2.8 million passengers. We will use this value as the "arrival/departure" volume for Novosibirsk.

The method of passenger traffic evaluation described above is not the only one. There exists a gravity model of trade [10], where passenger traffic between two nodes is proportional to the product of their total income and inversely proportional to the distance between the nodes. Moreover, the total time spent on transportation can be used as a minimizable indicator. In this case, we must account not only for the actual airborne time, but also for the time spent on commuting to the airport, registration, boarding, transfer between flights, and time spent commuting between the airport and railway station, etc. We can use combined cost, which includes transportation expenditures proper as well as cost of time spent traveling, which, naturally, depends on the income of citizens of a particular node.

3. Parameter variability

Forecasting the transportation system development is done by varying a range of expert-selected input parameters, including

- carrying capacity for the given means of transportation and selected shoulder or node;
- transportation cost (tariff) of the given product and means of transportation along the given shoulder or its processing in the given node.

Extending the system by implementing new types of variators which do not significantly extend the capabilities of the system but significantly simplify its use has turned out to be useful:

- coefficient for the means of transportation, by which all base tariffs connected with this means of transportation are multiplied;
- coefficient for the means of transportation by which all carrying capacities connected with this means of transportation are multiplied; and
- coefficient by which the passenger traffic increases in all incoming and outgoing directions from a given node.

For each variable parameter, an interval of acceptable values is defined. This interval usually contains the current value of the parameter, with the limits reflecting expert evaluations for the forecast period. The resulting multi-dimensional rectangle [7] is covered by a multitude of combinations of parameter values. MIX-PROSTOR allows two methods of such coverage:

- *incremental*, where each parameter changes within its interval with a specifiable increment, and
- *stochastic*, where the number of points randomly selected within a multidimensional rectangle is indicated.

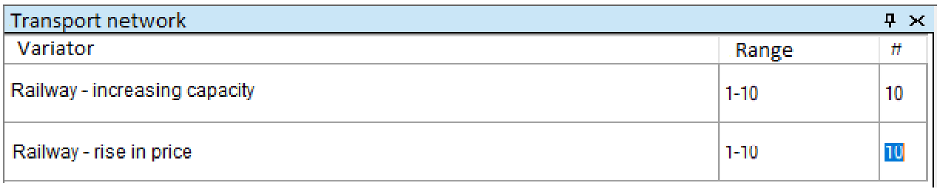
The stochastic method may prove more efficient from the applied point of view.

For each parameter combination, we look for a solution for a problem of linear programming, minimizing the total cost of transportation. Currently, MIX-PROSTOR uses the Google OR solver for this purpose [8].

The pool of resulting solutions may be very large, making expert analysis virtually impossible. For this reason, we clusterize solutions, comparing them by the similarity of their traffic streams along separate shoulders. From each cluster, a typical representative is selected, which is then presented to the expert for analysis.

4. Experiment and results

Below, we consider one of the performed experiments, studying the effect of railway transportation parameters on passenger traffic. First, we (roughly) define the coefficients of the carrying capacity change and tariffs for the whole railway network, as shown in Figure 1.


Transport network

Variator	Range	#
Railway - increasing capacity	1-10	10
Railway - rise in price	1-10	10

Figure 1. Setting parameter value intervals

This results in a total of 100 variants with up to a 10-fold increase both in the carrying capacity and in the railway tariff. Next, these variants are calculated, and the most typical transportation methods are selected from the pool of solutions.

The set of solutions divides into three typical methods of transportation, marked by red, yellow and green colors, which are distributed as shown in Figure 2.

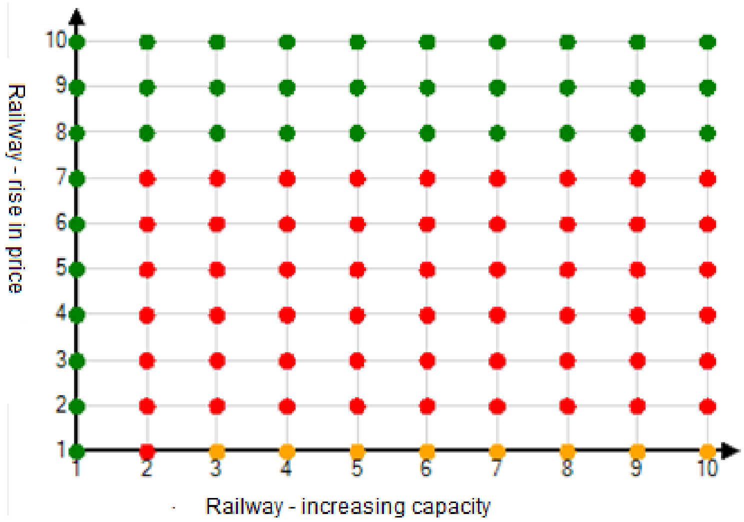


Figure 2. Rough estimate cluster partitioning

This information can be used as an estimate for a more precise definition of the variability intervals, for instance, by limiting the maximum coefficients to 3 and reducing the parameter increment by retaining the total number of solutions. This will result in a more detailed cluster partitioning (Figure 3).

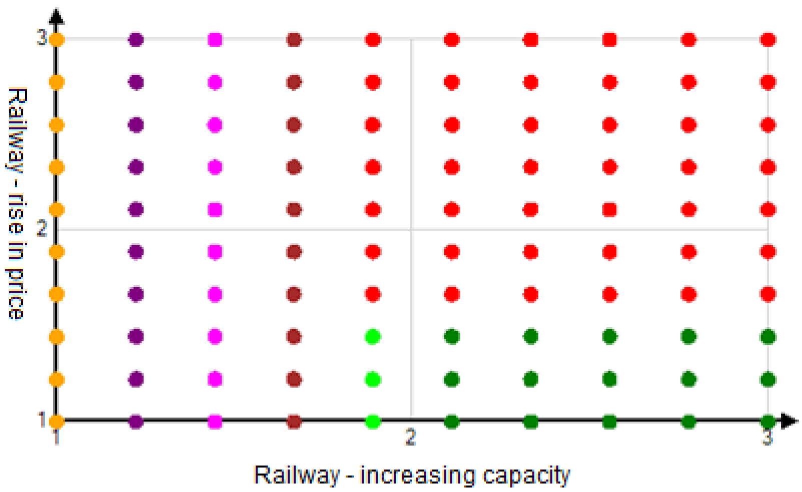


Figure 3. Cluster partitioning with accurate estimate

Next, the expert can analyze individual solutions. For instance, in the base state (both enlargement coefficients equal 1), which belongs to the yellow cluster, the transportation network looks as shown in Figure 4.

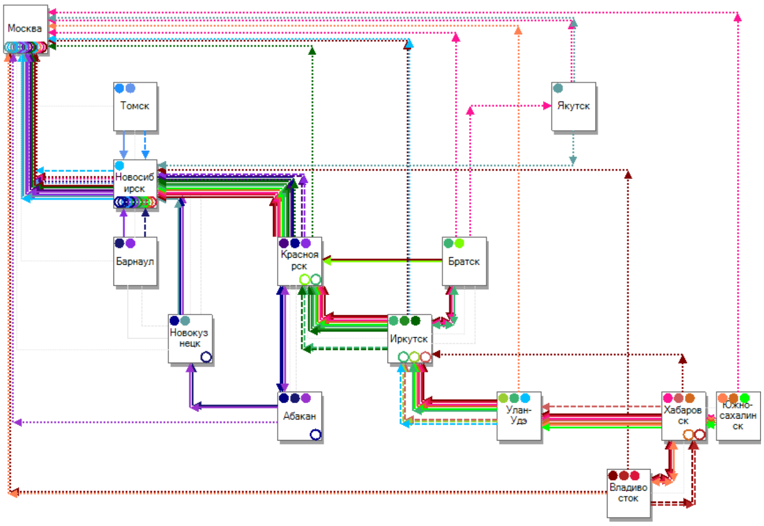


Figure 4. Base variant

Here, solid lines show railroad transportation, dash lines show highway transportation, and dot lines show air transportation. If the carrying capacity increases 1.4-fold (purple cluster), the use of air transportation decreases significantly, as shown in Figure 5.

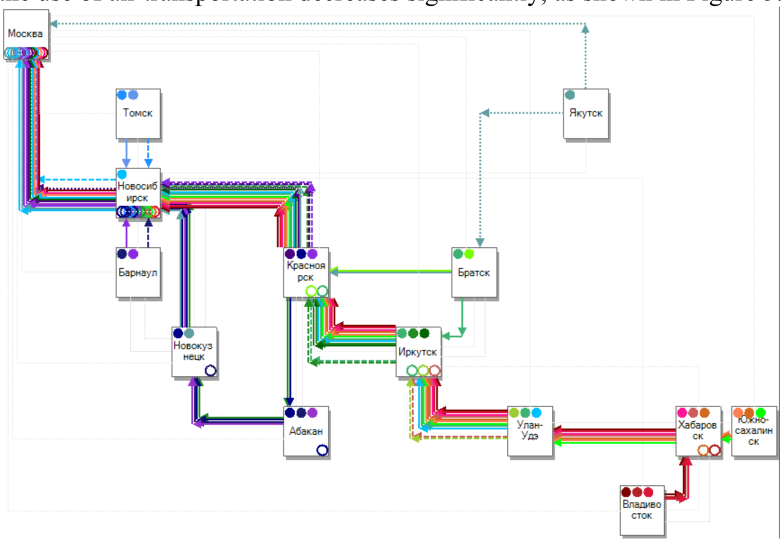


Figure 5. 1.4-fold increase of the railroad carrying capacity

The use of air transportation disappears almost completely if the railway carrying capacity is increased 1.6-fold. Yakutsk is an exception, as it has virtually no other connection except air transportation (Figure 6).

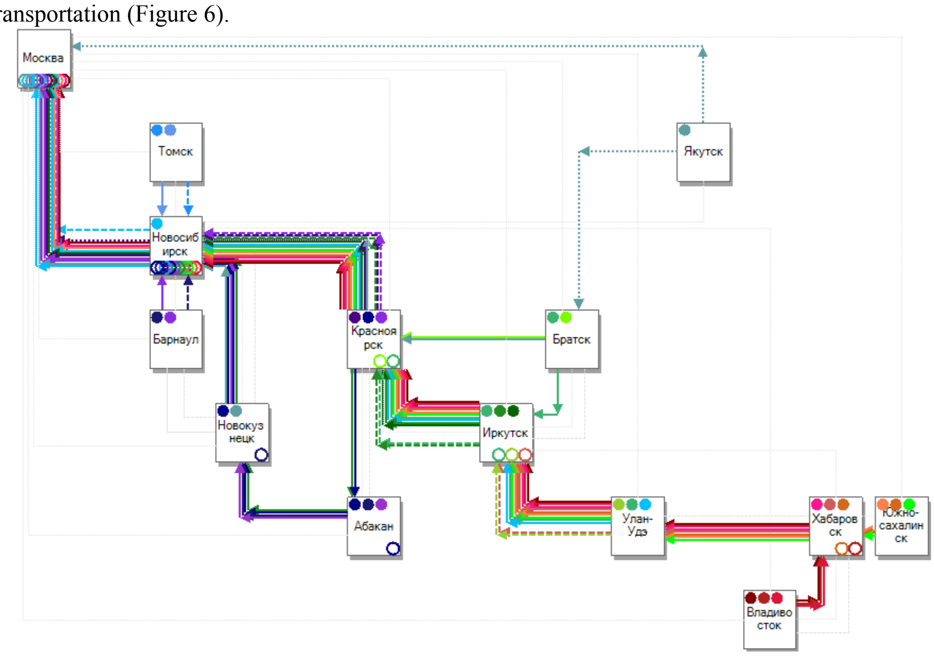


Figure 6. 1.6-fold increase of the railway carrying capacity

5. Transportation accessibility

Certain developments in the MIX-PROSTOR system concern transportation accessibility visualization. The average cost of passenger transportation (ACPT) for the pair <departure point, destination point> has been chosen as the defining indicator. We can easily demonstrate that for each generated solution it can be calculated as the average-weighted sum of the passenger transportation of this type along all transportation shoulders.

For a graphic representation of transportation accessibility and conurbations, we considered several variants with corresponding experiments. The first attempt consisted in using anamorphic maps, where the main concept is based on moving nodes in such a manner that the distance between them correspond with ACPT, preferably retaining their relative position in the process. Since ACPT is not a metric, and the triangle inequality does not hold true for it, such representation can only be approximate. To solve this problem, we used the power methods of graph placement; however, the results for the above-described goals were unsatisfactory.

The second attempt consisted in showing transportation accessibility from the point of view of an inhabitant of a specific node. The representation method consists in the following:

- the node of interest is placed in the center;

- all other nodes are placed relative to the first node in the direction corresponding to the real geographic direction;
- the distance to the adjacent nodes is proportional to the ACPT; the distance to other nodes is proportional to the length of the shortest route relative to the ACPT;
- the traffic volume is represented by the node size – the area of the corresponding circle is proportionate to the loading/unloading volume.

This works well when ACPT values are not too dispersed. Unfortunately, this is not the case here. In some experiments, traveling from Novosibirsk to Moscow costs 700 roubles, while traveling from Moscow to Yakutsk costs 60 000 roubles. As the result, in the visual representation, Novosibirsk is adjacent to Moscow and Yakutsk and is positioned somewhere on the other side of the diagram.

To solve this problem, we used the fisheye view technology, where distances are increased in the middle and shrunk along the edges. There exist multiple ways to achieve this. We used the method described in [9], where the distance d is transformed using the following formula:

$$T(d) = d_{max} * \frac{dist + 1}{dist + \frac{d_{max}}{d}},$$

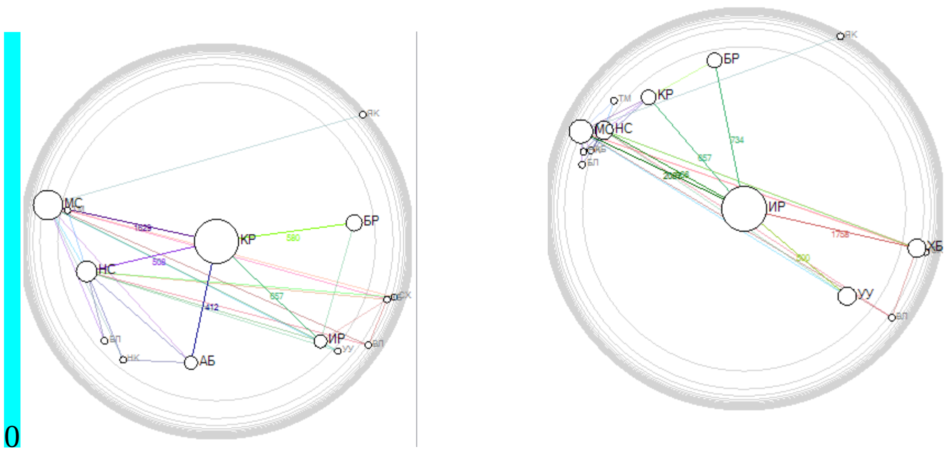
where

d_{max} is maximum distance between nodes;

dist is the distortion coefficient;

dist = 0 produces an identical representation, i.e., the image remains undistorted. We provide the possibility to vary *dist* interactively.

The concentric circles in Fig. 7 represent the ACPT and are equally spaced in reality. This means that for an inhabitant of the central node, everything in the gray zone is very far and very expensive.



(a) (b)

Figure 7. Representation of transportation accessibility: (a) Krasnoyarsk, (b) Irkutsk

Such representation may be useful for the expert evaluation of conurbations from the point of view of transportation accessibility. For instance, Figure 7(b) shows a distinct conurbation of Irkutsk – Bratsk – Ulan-Ude. On the other hand, Krasnoyarsk belongs both to the Irkutsk and Novosibirsk conurbations.

The MIX-PROSTOR system allows predicting changes in conurbation composition from varying the transportation system parameters. For instance, in the experiment described below, we simultaneously varied costs and railway carrying capacity. Fig. 8. shows the results of the experiment with Irkutsk selected as the central node.

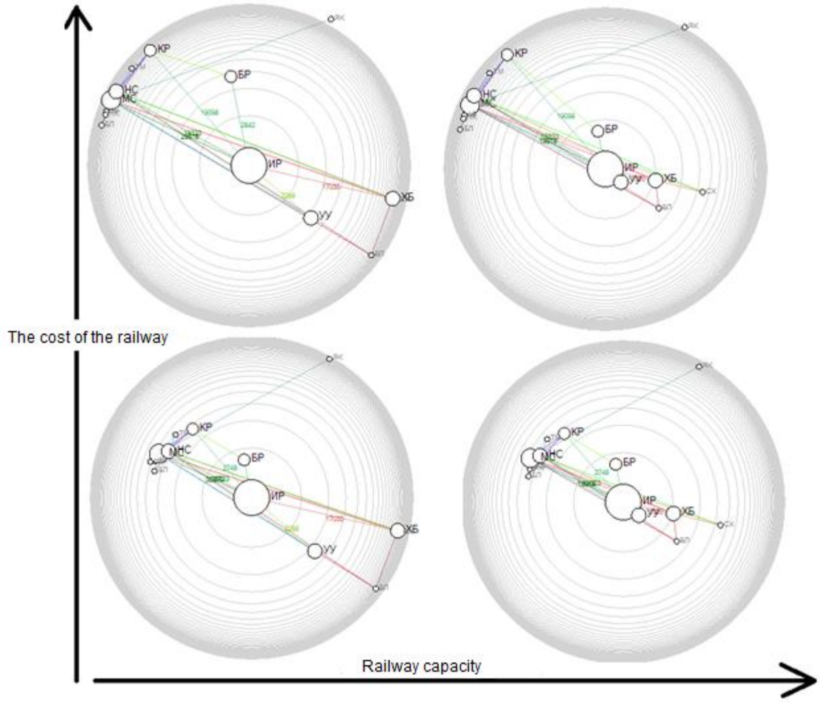


Figure 8. Conurbation change forecast

The analysis of the experiment results allows us to draw the following conclusions:

The upper left image (low carrying capacity – high price) effectively isolates the conurbation from the rest of the country.

The upper right (high carrying capacity – high price) approximates it to the Far East.

The lower left (low carrying capacity – low price) moves the conurbation closer to European Russia.

The optimal variant in the lower right image brings the whole country together (with the exception of Yakutsk).

Conclusion

This paper examines some problems of forecasting the development of the backbone transportation network in Asian Russia concerning passenger transportation. The MIX-PROSTOR system has proven to be a powerful tool that allows its users to identify and focus the attention of experts on the most relevant aspects [3, 6, 7]. Further research can be connected both with the integrated studies of passenger and cargo transportation and with extending the transportation network and expanding the range of the products transported.

References

[1] Problem regions of resource type: Asian Russia / Ed. V. A. Lamin, V. Yu. Malov. – Novosibirsk: SB RAS Publishing, 2005 (In Russian).

[2] Asian Russia: modeling economic growth in context of historical experience / V. A. Lamin, V. Yu. Malov. SB RAS: IEIPO, Institute of History, V. B. Sochava Institute of Geography, L. A. Melentyev Institute of Energetics, ISDCP. – Novosibirsk: SB RAS Publishing, 2012. – (Integration Projects; Issue 34) (In Russian).

[3] Situation room as element of organization of expert community: problems of planning and forecasting / Ed. G. A. Untura. – Novosibirsk: IEIPO SB RAS, 2018 (In Russian).

[4] Bulyonkov M. A., Karpan V. V., Malov V. Yu., Marusin V. V., Radchenko V. V. Conceptual problems of building a Model-Information-Cartographic System (MIX) // Modeling of Industrial and Regional Systems based on GIS and Information Technologies / Ed. Yu. Sh. Blam, V. V. Radchenko. – Novosibrisk: IEIPO SB RAS, 2011. – P. 5-28 (In Russian).

[5] Vorobyeva V. V., Malov V. Yu., Radchenko V. V., Potter M. V., Serebryannikov I. E. Model for forecasting of backbone transportation network development // Modeling of industrial and regional system based on GIS and information technologies / Ed. Yu. Sh. Blam, V. V. Radchenko. – Novosibrisk: IEIPO SB RAS, 2011. – P. 68-96 (In Russian).

[6] Bulyonkov M. A., Filatkina N. N. Situation analysis in the MIX-PROSTOR transportation forecast system // Informatsyonnye Tekhnologii. – 2013. – No.8. – P. 43-52 (In Russian).

[7] Bulyonkov M. A., Nesterenko T. V. Automation of research of backbone transportation network // SibSUTI Bulletin. – 2019. – No.47. – P. 45-54 (In Russian).

[8] Google OR-Tools. – https://developers.google.com/optimization/ (accessed on 03.03.2022).

[9] Sarkar, Manojit, Brown, Marc Graphical Fisheye Views of Graphs. // Proc. of the ACM CHI 92 Human Factors in Computing Systems Conference June 3-7, 1992, Monterey, California / Ed. Bauersfeld, Penny, Bennett, John, Lynch, Gene. – P. 83-91. – https://www.acm.org/pubs/articles/proceedings/chi/142750/p83-sarkar/p83-sarkar.pdf.

[10] Carrère, Céline; Mrázová, Monika; Neary, J. Peter Gravity without Apology: The Science of Elasticities, Distance, and Trade // The Economic Journal [Electronic resource] . – 2020. –Vol. 130, No. 628. – P. 880–910.